

1.0 INTRODUCTION

1.1. Background

The United States of America Federal Communications Commission's (FCC) Advisory Committee on Advanced Television Systems (ACATS) was formed in 1987 to advise the FCC regarding the adoption of a new, advanced broadcast terrestrial television service (ATV). Systems Subcommittee Working Party 2 (SSWP2), Test and Evaluation, was given the task of developing and carrying out test procedures. In May 1993 the Grand Alliance (GA), a consortium of the formerly competing HDTV system proponents, was formed to combine the best efforts of each.

The Audio Specialist Group of the Technical Oversight Group (TOG) of the Grand Alliance was given the job of testing the multichannel coders under consideration. These coders supply five full-bandwidth channels (left, right, center and two surround channels), plus an additional, very low frequency subwoofer enhancement channel. Tests were carried out in July and December of 1993 in a listening room at the THX offices at Lucas Films' Skywalker Ranch in California. The results of these tests led the GA to select the AC-3 coder. In November of 1994 the AC-3 codec was adopted as a standard by the U. S. Advanced Television Systems Committee (ATSC). This standard is documented in ATSC document A/52.

ACATS required that the complete integrated GA HDTV system undergo comprehensive testing of actual operating hardware. SSWP2 decided that these tests should include subjective audio tests.

The goals of the audio test were determined to be:

- ~ Verify that the performance of the fully integrated AC-3 coder is as good as, or better than, the original 1993 coder.
- ~ Measure the basic audio quality of the AC-3 coder operating in the 5.1 channel mode at 384 kb/s with 5.1 channel reproduction.
- ~ Measure the basic audio quality of the AC-3 coder operating in the 5.1 channel mode at 384 kb/s with 2 channel reproduction.
- ~ Measure the basic audio quality of the AC-3 coder operating in the 2 channel mode at 256 kb/s with 2 channel reproduction.

1.2 Test Plan

The test plan developed by SSWP2's Audio Task Force to satisfy these goals is based entirely on ITU-R Recommendation BS. 1116, "Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems". Three types of listening sessions were conducted:

- ~ 5.1 channel source, 5.1 channel reproduction.
Final GA AC-3 coder and the 1993 coder, both at 384 kb/s
- ~ 5.1 channel source, 2 channel reproduction.
Final GA AC-3 coder at 384 kb/s.

~ 2 channel source, 2 channel reproduction.
Final GA AC-3 coder at 256 kb/s.

2.0 LISTENING PANELS

2.1 *Expert Listeners*

Twenty-two (22) "expert" listeners, i.e. those people who have extensive listening experience, especially, wherever possible, experience listening for artifacts of perceptual coders, were recruited by issuing a "call for expert listeners" to various organizations such as the Audio Engineering Society local Washington DC area chapters as well as broadcast and cable organizations.

Their ages ranged from 22 to 61 years, there were more men than women (there were 5 women) and most were audio mixers/recording engineers or audio engineers involved in R & D, and frequently they were involved in music as a hobby in some capacity.

2.2 *Critical Test Material Selection Panel*

A small panel of special expert listeners, all of whom are involved in psychoacoustic endeavors by profession, was chosen to reduce the number of possible test material selections to ten for each test type. Three members of the Audio Task Force also attended and contributed, but did not attempt to influence decisions regarding the choices.

The ten best critical selections in each case were identified by repeated review and note taking. These selections and the random orders were then taken to an audio work station where randomized tapes for the expert-listening double-blind tests were prepared.

It should be noted that when the new-coder tape was played, just after the old-coder tape, a clear improvement was noticed immediately by the panel. It continued to appear to the experimenter that the attempted identification, comments and remarks being made were still rather random, as opposed to correct, on this new-coder tape.

Further details about the critical listening panel and its task, listening room, equipment, order of presentation etc. can be found in Annex B and in Annex C.

3.0 PROGRAM MATERIAL

The audio test plan called for audio sequences to be passed through the entire GA system, from audio encoder, through system multiplexing and transmission modulation, demodulation and demultiplexing, and finally audio decoding. It was therefore necessary to pre-select a set of audio sequences to be assembled on a tape which would be put through the entire system. Ten sequences were to be used for each final listening test. In order to allow for a post-selection of the most critical sequences (required by BS.1116), it was decided to pre-select 20 sequences to be passed through the coder.

Two sets of test material were collected: 2-channel stereo, and 5.1-channel multichannel. The universe of multichannel sequences was the set of 10 sequences previously prepared by SSWP2, the 64 selections from the 1994 MPEG multichannel testing, and 6 new selections obtained from a variety of sources. Included in the SSWP2-provided sequences were the

selections which had been specifically produced for this purpose in Salt Lake City early in 1993 (including the now well-known, "Glockenspiel and Timpani"), as well as a number of film sound excerpts. A pre-selection process conducted by the SSWP2 Audio Task Force in Nov. '94 picked 20 of these sequences. The pre-selection of the 20 sequences was done by considering the demonstrated (by the July '93 GA tests and the '94 MPEG tests) criticality of some sequences, as well as a desire to create a grouping somewhat representative of a wide variety of material.

The universe of 2-channel stereo material was composed of the selections (essentially identical to those used during one set of MPEG-1 audio tests) from previous 2-channel tests conducted by ACATS as part of the initial ATV system testing two years earlier (when the 5 initial HDTV proponent systems were tested), as well as the sequences used during EIA/DAR tests at the CRC in 1994 and 1995.

3.1 Selection of Final Critical Test Material

A list of the final test material selections is attached as Annex A. The material comes from film, a Salt Lake City 5.1 channel recording session, a few experimental professional 5-channel recordings, the EBU SQAM disc (stereo only) and from the MPEG test material from their 1994 listening test material. It included the harpsichord piece which was the most difficult for all the coders in the MPEG listening tests, as well as four other selections which had been chosen as "critical" through the coders competing in that forum. The most critical five selections (passed through the 1993 coder) were used for training as well.

3.2 Randomization of Presentation Order

A Digi Design Pro Tools Macintosh-based work station was used to make the 14 randomized digital tapes recorded onto super VHS video tape using an Alesis ADAT machine. It was also used to add a 200 ms linear fade at the end of two of the 5.1 channel selections to avoid abrupt-sounding endings.

4.0 LISTENING CONDITIONS AND TEST EQUIPMENT

Specific requirements for listening conditions (to ensure comparable test results) from Rec. BS. 1116 were followed. The July '93 GA listening tests had been conducted in a room which met both the BS.1116 specifications and the THX specifications. The rooms offered for both the selection process in California and the expert listening in Washington DC also met THX specifications and generally met the requirements of BS. 1116. Details regarding the listening rooms can be found in Annex C.

The maximum SPL of the 5.1-channel test selections reached 78 to 90 dB and that of stereo reached 85 to 93 dB SPL A weighted.

A light box in the front of the room illuminated one of three panels, "Ref" or "A" or "B" one second prior to hearing the selection, the illumination was turned off at the selection's end.

5.0 EXPERIMENTAL DESIGN

Per ITU-R Recommendation BS.1116, the "triple stimulus with hidden reference" test method was employed. Tape recordings were prepared in several random presentation orders, using the recommended "Ref" "A" "B" order. The use of the tape-base test method allowed more than one listener to be tested in any given test session, improving the

efficiency of the test. The subjects were run in groups of two to five and were allowed to sit in any of eight possible listening locations. Each heard every multichannel test selection in both possible orders, (coder before reference and the reverse) twice. Audiometric tests were not performed.

5.1 Test Method

The "triple stimulus, hidden reference, double blind" test method was employed. Two identical triplets were presented, "Ref," -2s- "A," -2s- "B," -2s- "Ref," -2s- "A" -2s- "B" ----15s----, then the subject would score "A" and "B" by placing a mark on a continuous "thermometer-like" scale or by writing a number on the 5-grade impairment scale to one decimal place. The ITU-R (formerly CCIR) Impairment Scale which was used, incorporates the descriptors:

5. Imperceptible
4. Perceptible, but not Annoying
3. Slightly Annoying
2. Annoying
1. Very Annoying

5.1.1 Procedure

The listeners and experimenter met for about twenty minutes before the start of training, to review the test instructions (see Annex D), a confidentiality statement, a table of possible artifacts to listen for (MPEG ISO/IEC JTC1/SC29/WG11 1994 Report, Appendix C, Annex C, pg. 5) and the general schedule for the day.

An hour was then taken to listen to the training tape a minimum of two times and to discuss the format of the test method, answer sheet and numbering or marking system to be used.

The training tape contained the five most sensitive of the ten test-material selections in three pairs: "Ref" -2s- "test", -2s- "Ref" -2s- "test", -2s- "Ref" -2s- "test" ----15s----.

5.1.2 Randomization and Order of Presentation

There were two blocked pseudo-random orders of the stereo 2/2 test material, each about 15 minutes long and four random orders of each of the other test conditions (i.e. stereo mixdown, 5.1-channel original coder and 5.1-channel new coder), each about one-half hour long. The designation "pseudo" is due to the imposition of constraints or rules upon true randomness such as the rule that any selection will not follow the same selection or itself more than once. Stereo tapes were kept in pairs to help the listeners, but their order was alternated each day (for instance, tapes 9, 10, 11, and 12 were stereo mixdown tapes (5.1 to 2 channels) all contained the same material, but recorded in different, randomized presentation orders: tapes 9 then 10 were presented on day-one, tapes 11 then 12 on day-two, 10 then 9 on day-three, 12 then 11 on day-four and so on, using a different daily presentation order of eight (of 14 total tapes) total daily tapes each day.

The test design provided data from two presentations of every test selection, in both possible orders, (A before B and B before A were on different tapes), resulting in two data points from every listener on every sequence.

6.0 STATISTICAL ANALYSIS AND RESULTS

One goal of this listening test was to ensure that the new coder was as good as or better than the 1993 coder. An analysis of variance (ANOVA) shows that the Harpsichord was correctly detected in the 1993 coder but that no selection was correctly identified on the new coder. Figures 1, 2 and 3 show this effect graphically at 95% and 99% confidence levels. The detailed results are shown in Tables 1 and 2.

Figure 1

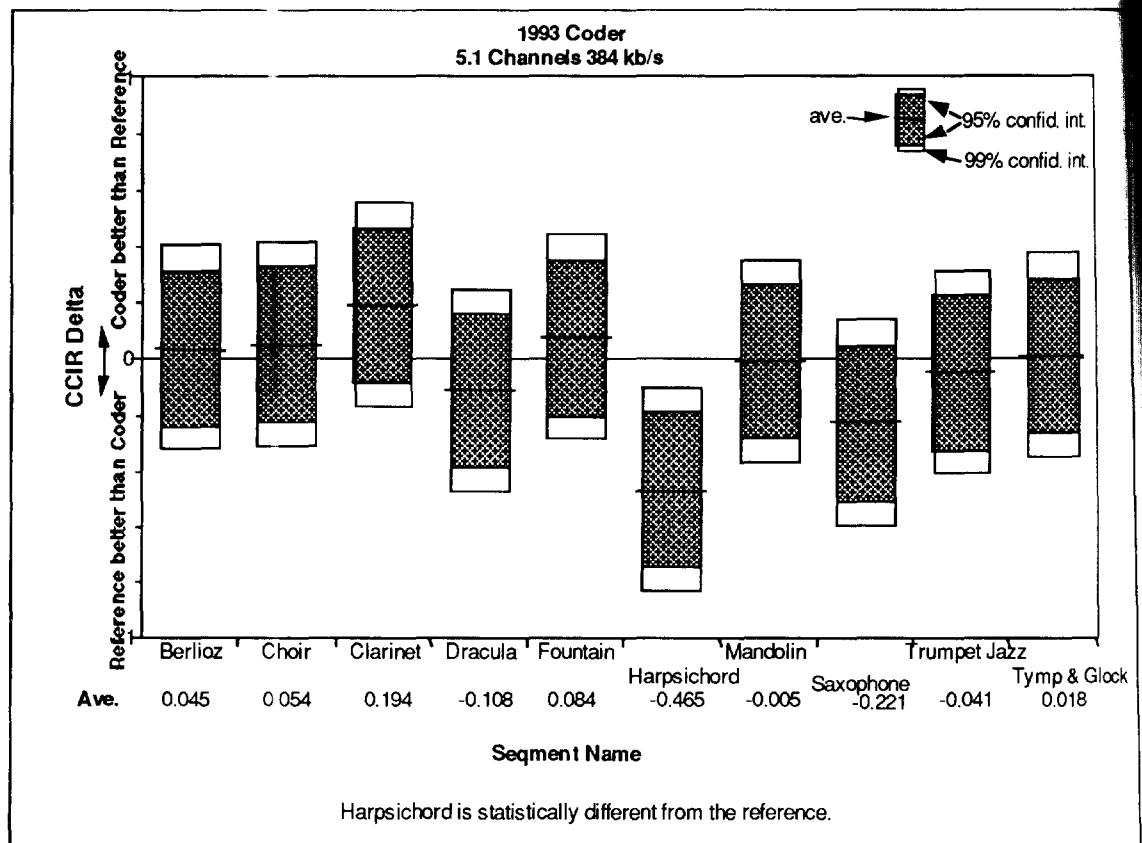


Figure 2

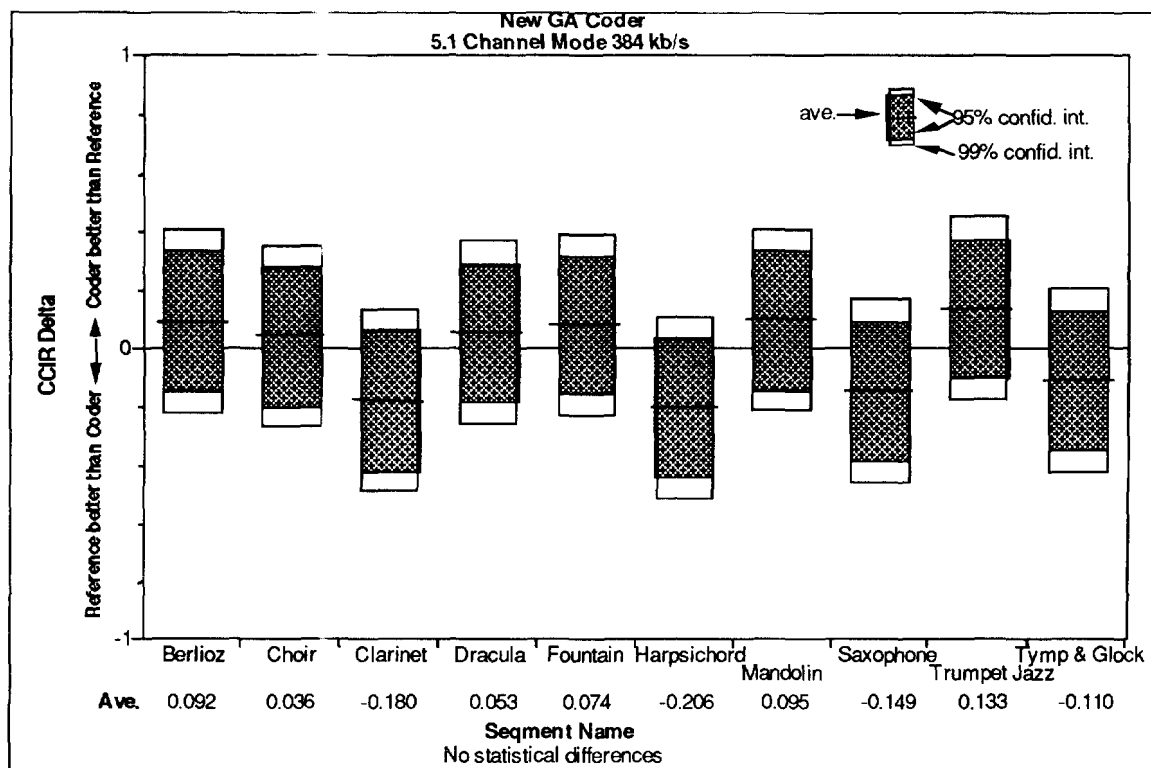


Figure 3

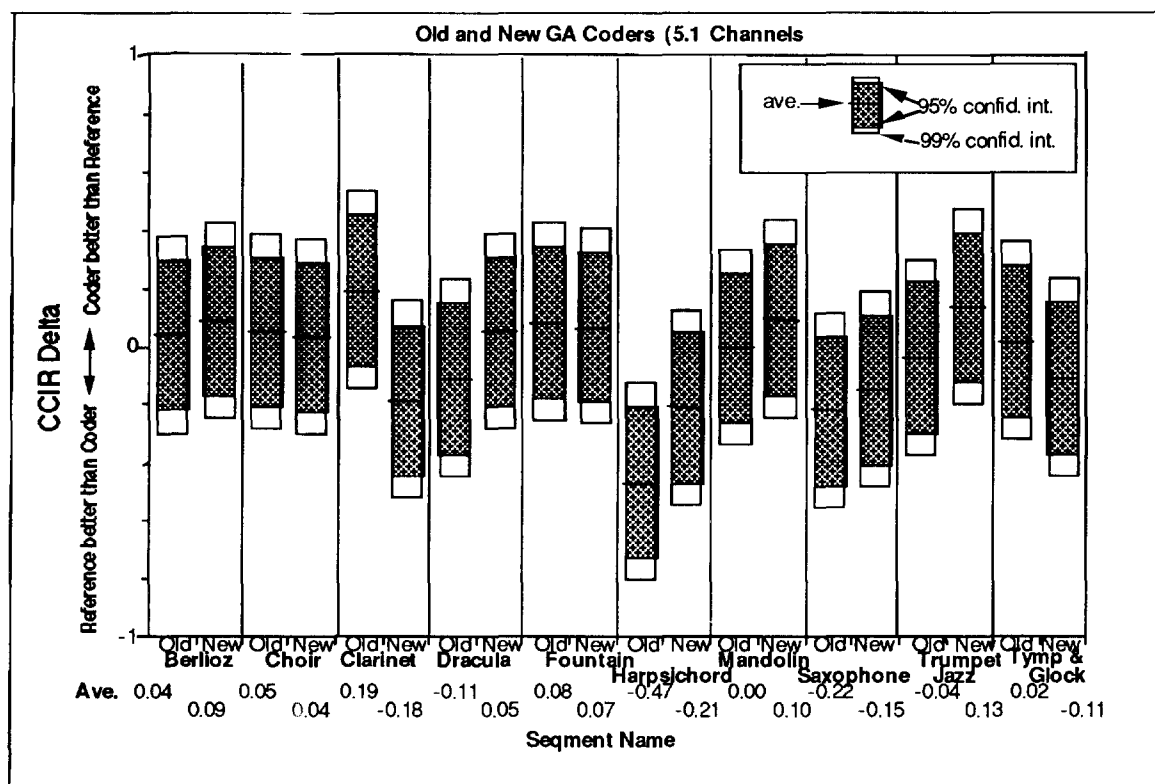


Table 1 July '93 AC-3 Coder 3/2/.1 mode, 384 kb/s (n=44)

Segment	Average	Std Deviation	95% Con. Int	Prob t	Prob signed-rank
Berlioz	0.044	0.951	-0.24, 0.33	.621	.565
Choir	0.054	0.754	-0.18, 0.28	.682	.447
Clarinet	0.194	0.889	-0.08, 0.46	.923	.921
Dracula	-0.108	0.663	-0.31, 0.09	.144	.185
Fountain	0.084	0.557	-0.09, 0.26	.829	.800
Harpsichord	-0.465	1.348	-0.87, -0.06	.013	.019
Mandolin	-0.005	0.855	-0.26, 0.26	.49	.45
Saxophone	-0.224	0.907	-0.50, 0.05	.056	.025
Trumpet Jazz	-0.041	0.837	0.30, 0.21	.37	.507
Tymp&Glock	0.02	1.19	-0.34, 0.38	.54	.517

Harpsichord is statistically different from the reference. Saxophone is not significant at 5% (.05) for the t test, but is significant using the signed rank test (.025). It is useful to point out that a 95% confidence level in this case is equivalent to a 97.5% level since the coder was not allowed to be better than the reference.

Table 2 GA AC-3 Coder 3/2/.1 mode, 384 kb/s (n=44)

Segment	Average	Std. Deviation	95% Con. Int	Prob t	Prob. signed-rank
Berlioz	0.09	1.05	-0.23, 0.41	.718	.737
Choir	0.04	0.591	-0.14, 0.22	.658	.440
Clarinet	-0.18	0.680	-0.39, 0.03	.043	.069
Dracula	0.05	0.629	-0.14, 0.24	.709	.796
Fountain	0.07	0.670	-0.13, 0.28	.766	.839
Harpsichord	-0.21	0.957	-0.50, 0.08	.080	.123
Mandolin	0.10	0.734	-0.13, 0.32	.803	.632
Saxophone	-0.15	0.716	-0.37, 0.07	.089	.190
Trumpet Jazz	0.13	0.834	-0.12, 0.39	.851	.889
Timp & Glock	-0.11	0.998	-0.41, 0.19	.235	.24

Clarinet is statistically different using the t test (.043), but is not by the signed-rank test.

The other goal of these listening tests was to evaluate the basic audio quality of the new 5.1-channel coder. No statistically reliable correct detections took place. Since the coder could not be identified, the conclusion must be that the new multichannel coder is indistinguishable from the source (i.e. transparent) under these conditions. In comparing the results of the old versus the new coder, it should also be noted that the magnitude of all differences is less on the new coder (this comparison is valid since it is made on the same test, with the same subjects, same method etc.).

The stereo mixdown test (Figure 4) showed a statistically significant correct detection on the harpsichord and perhaps on God Be With You. See Table 3.

Figure 4

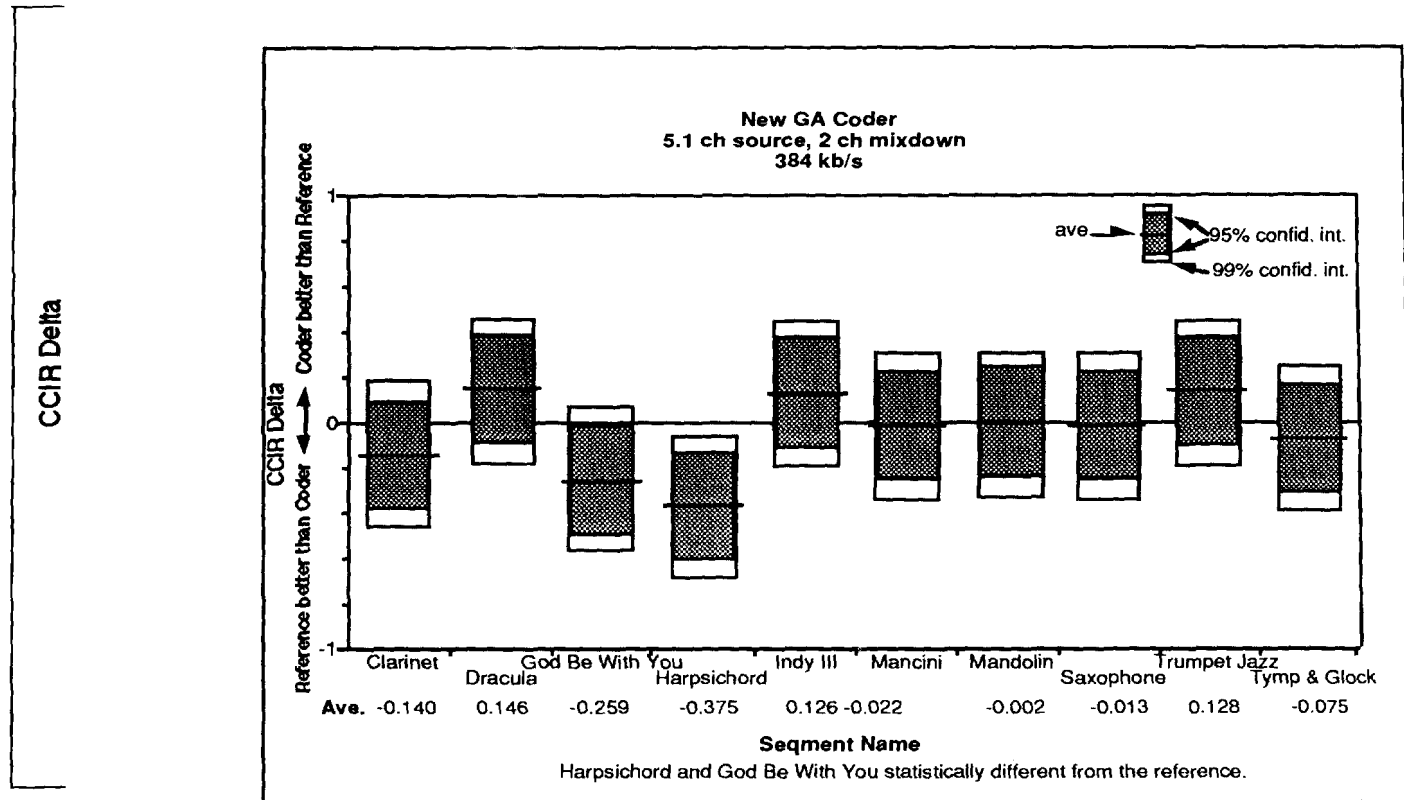


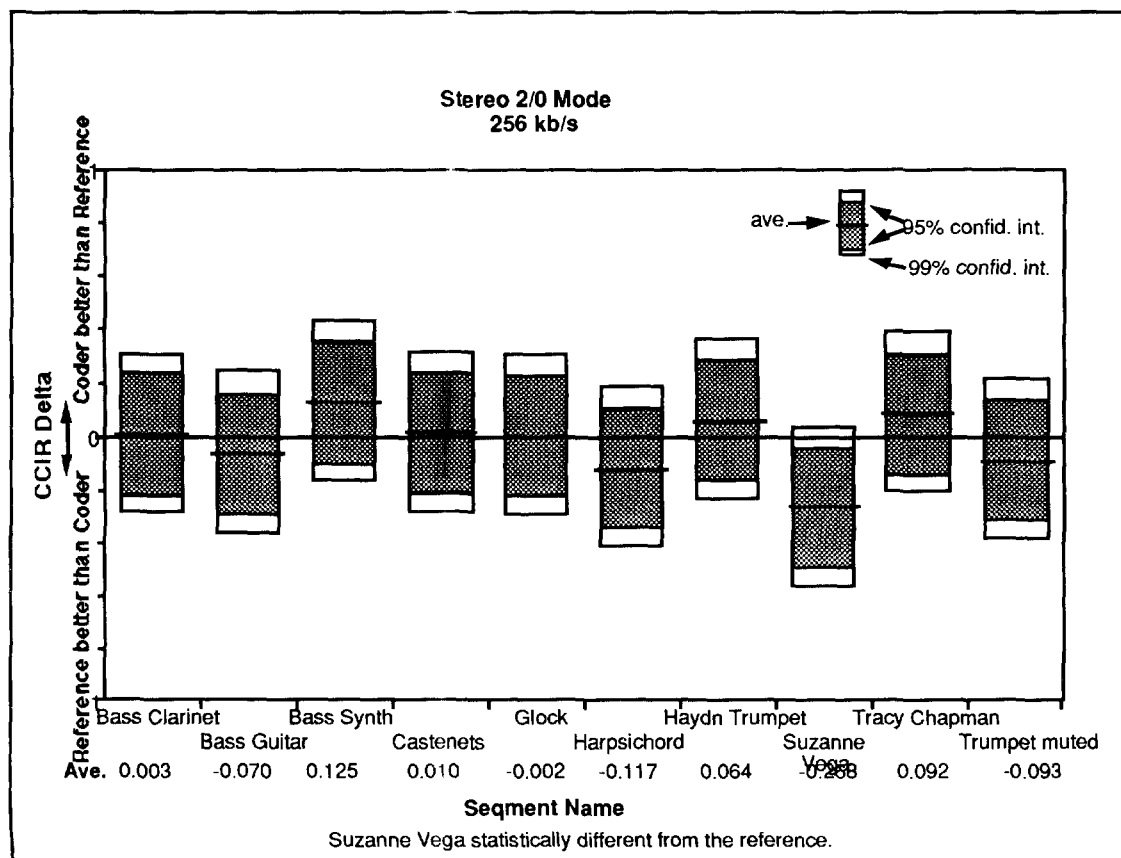
Table 3 2 ch Stereo Mixdown of GA AC-3 Coder, 5.1 ch mode, 384 kb/s (n=44)

Segment	Average	Std. Deviation	95% Con. Int	Prob t	Prob. signed-rank
Clarinet	-0.14	1.02	-0.45, 0.17	.184	.058
Dracula	0.15	0.684	-0.06, 0.35	.918	.992
God Be With You	-0.26	0.937	-0.54, 0.03	0.037	.097
Harpsichord	-0.37	0.993	-0.66, -0.09	.005	.006
Indy	0.13	0.720	-0.09, 0.34	.874	.878
Mancini	-0.02	0.692	-0.23, 0.19	.417	.319
Mandolin	-0.002	0.775	-0.24, 0.23	.495	.500
Saxophone	-0.01	0.670	-0.23, 0.20	.452	.530
Trumpet Jazz	0.13	0.641	-0.07, 0.32	.904	.863
Tyrmpe&Glock	-0.08	0.842	-0.33, 0.18	.278	.320

Harpsichord is statistically different. God Be With You is significant by the t test, but not by the signed-rank test.

The stereo 2-channel test (Figure 5 and Table 4) showed a statistically significant detection only on Suzanne Vega at the 95% level. However, if one looks at the data from only the seven best listeners (i.e. \geq or $> 75\%$ accuracy) or at the 99% level, that difference disappears.

Figure 5



**Table 4 GA AC-3 Coder 2/0 mode, 256 kb/s
All Listeners (n=44)**

Segment	Average	Std. Deviation	95% Con. Int	Prob. t	Prob. signed- rank
Bass Clarinet	0.003	0.837	-0.25, 0.26	.510	.394
Bass Guitar	-0.07	0.70	-0.28, 0.14	.254	.475
Bass Synth	0.12	0.724	-0.10, 0.34	.870	.707
Castanets	0.01	0.416	-0.11, 0.14	.566	.483
Glock	-0.002	1.022	-0.31, 0.31	.495	.581
Harpsichord	-0.12	0.868	-0.38, 0.15	.187	.135
Haydn Trumpet	0.06	0.643	-0.13, 0.26	.752	.744
Suzanne Vega	-0.27	0.673	-0.47, -0.07	.006	.002
Tracy Chapman	0.09	0.776	-0.14, 0.33	.783	.853
Trumpet muted	-0.09	0.906	-0.37, 0.18	.249	.399

The only segment different from the reference is Suzanne Vega.

**Table 5 GA AC-3 Coder 2/0 mode, 256 kb/s
Listeners who identified the coder correctly $\geq 75\%$ (n=14)**

Segment	Average	Std. Deviation	95% Con. Int.	Prop. t	Prob. signed- rank
Bass Clarinet	0.06	0.875	-0.45, 0.56	.594	.594
Bass Guitar	-0.18	0.774	-0.63, 0.27	.202	.375
Bass Synth	0.20	0.572	-0.13, 0.53	.889	.812
Castanets	-0.07	0.267	-0.23, 0.08	.168	.500
Glock	-0.40	0.952	-0.95, 0.15	.068	.081
Harpsichord	-0.19	0.598	-0.54, 0.15	.123	.125
Haydn Trumpet	-0.06	0.458	-0.33, 0.20	.309	.250
Suzanne Vega	0.02	0.753	-0.41, 0.46	.540	.447
Tracy Chapman	0.30	0.618	-0.05, 0.66	.995	.953
Trumpet muted	-0.036	0.745	-0.79, 0.07	.048	.125

No significant differences.

7.0 CONCLUSIONS

It can be concluded that :

- 1) the audio quality of the fully integrated GA coder is better than that of the original 1993 coder.
- 2) the audio quality of the GA coder in the multichannel mode was indistinguishable from that of the source.
- 3) the audio quality of the GA coder in the 5.1 mode with 2 channel reproduction, while it can be detected by some expert listeners on some audio test material, is very nearly transparent, better than grade 4.5 on the 5-point impairment scale.
- 4) the audio quality of the GA coder in the 2 channel mode, while it may have been detected on the Suzanne Vega sequence, this was probably a statistical fluke (at the 95% level, oddities will show up one in 20 times) it is very nearly transparent (better than grade 4.7 on the 5 point impairment scale).

REFERENCES

"Report on the MPEG/Audio Multichannel Formal Subjective Listening Tests". Draft Report from ISO/IEC JTC1/SC29/WG11 NO686, March 1994.

ITU-R Recommendation BS. 1116 "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems", March 1994.

"Digital Audio Compression (AC-3), U.S. Advanced Television Systems Committee, Doc. A/52, November 1994.

Segme
Bass
Bass
Bass
Casta
Gloc
Harp
Hay
Suz
Trac
Tru

The

Se
Ba
Ba
B
C
C
F
F
S
,
.

ANNEX A

Audio Test Material: 5.1 channel

1) Fountain (water & classical piano, Mozart)	MPEG test material
2) Sax (multiple instrument, jazz)	MPEG test material
3) Harpsichord (single voice)*	MPEG test material
4) Clarinet & rain (broken chord)**	MPEG test material
5) Jazz (trumpet and piano)	MPEG test material
6) Berlioz II (woodwind, stage music)	MPEG test material
7) Timpani & Glockenspiel	Salt Lake City Quality test material
8) Choir (acapella, 4 voices)	Harmonia Mundi
9) Dracula Music Box (f. voice & music box)	Columbia Pictures
10) Fife & Drum (w/ mandolin)	Windam Hill

*this selection was most difficult for all coders in the MPEG tests

** this selection was an MPEG training sequence.

(Fountain, Mancini, Indiana Jones and Harpsichord were "critical" listening test material through the MPEG coders.)

Audio Test Material: 5.1 to 2.0 channel mixdown

1) Mancini (orchestra, violins, brass & percussion)	MPEG test material
2) Indiana Jones III (voice w/ orchestra)	Lucas Films
3) God Be With You (Choir, applause & ambiance)	Salt Lake City test material

(these three replaced Fountain, Berlioz and Choir above and the rest remained)

Page

Audio Test Material: 2 channel

REF

"Rep
Repo

ITU-
impa

"Digi
A/52

- | | |
|---------------------|-------------------|
| 1) Suzanne Vega | A&M 395 136-2 |
| 2) Tracy Chapman | Elektra 960 774-2 |
| 3) Bass Synthesizer | Swedish Radio |
| 4) Castanets | EBU SQAM disc |
| 5) Bass Guitar | Swedish Radio |
| 6) Haydn trumpet | EBU SQAM disc |
| 7) Bass Clarinet | EBU SQAM disc |
| 8) Glockenspiel | EBU SQAM disc |
| 9) Harpsichord | EBU SQAM disc |
| 10) Trumpet | EBU SQAM disc |

ANNEX B

Selection Panel Report

The selection panel for the ACATS SSWP2 Audio Task Force test material met at Snell Technology Center in Valley Village, CA May 1 - 3, 1995. Members of the selection panel were found through professional contacts. They numbered n=5 the first day and n=4 the second day. All had a great deal of experience in controlled listening tests, and some specific training in potentially audible problems with perceptual codices. Listening was done by hearing first an uncoded reference sequence several times and then the coded samples of the same excerpt (i.e. the listening order they chose was Ref., Ref., Coder, Ref., Coder ... [Ref. coder repeated if desired]). They did no "blind" listening and they were encouraged to discuss what they thought they heard and to help each other. They also were encouraged to change seats so as to listen from many different listening positions. Twenty samples of program material were provided, obtained from original recordings, or from multichannel film sources described elsewhere. The job of the panel was to find 10 samples which had the greatest sensitivity to codec errors.

The panel started with the original 1993 multichannel codec, comparing it to the known reference. The panel, with good unanimity, believed that certain of the pieces of program material did show up differences between the codec and the reference. Those items of material were noted for later re-listening. Next, the new GA multichannel codec was compared with the reference. Again the panel agreed, with only shades of the same opinion revealed in the discussions, so good unanimity was achieved. The opinion was that this codec was much harder to distinguish from the reference, with a smaller number of items on which any difference was audible compared to the older codec. Also, the difference between coded and uncoded signal (if any) were smaller than on the older codec. All of the items were auditioned on both the 1993 and the new GA codec.

Several of the participants reported that the 2 channel tests were even less revealing than the multichannel case.

The reason that the 1993 original coder was first was that it was known from experience that correct detections might take place on the "Glockenspiel and Timpani" and "Harpichord" selections. This turned out to be true. A definite and immediate improvement was noted by the whole group when those selections were played on the new GA coder; there was a clear difference. It was not clear, with this panel, whether any correct detections subsequently took place on the new coder.

The stereo selections were also difficult to select and again it was not apparent to the experimenter whether correct detections were taking place (expert listeners almost always comment on differences, real or imagined).

It should also be noted that in previous listening tests at the BBC and Deutsche Telekom the following selections have been determined to be critical: "Timpani and Glockenspiel", "Harpichord", "Indiana Jones", "Berlioz", "Clarinet and Rain", "Rock " and "Fountain". Both MPEG locations found clear artifacts in "Harpichord" and "Mancini"; "Timpani and Glockenspiel" and "Clarinet and Rain" were used as training sequences because they "evoked a number of artefacts that are characteristic of various psychoacoustic based

coders"¹ with "Timpani and Glockenspiel" ranked first of four, and "Clarinet and Rain" ranked third of four training sequences.

The experts who selected the final 10 selections for these tests agreed that "Timpani and Glockenspiel", "Fountain", "Trumpet Jazz" and "Harpsichord" exhibited clear artifacts on the '93 encoder. Several of the remaining selections were accompanied by comments from the experts as "minor difference, hard to hear, not very noticeable". The new coder had comments like "subtle, minor, slight, no audible difference" including "Timpani and Glockenspiel" (very minor) and "Harpsichord" (no audible difference). The experts also commented on the artifacts being difficult to hear for the stereo selections.

¹Annex C page 3, Report on the MPEG/Audio Multichannel Formal Subjective Listening Tests, March 1994, ISO/IEC JTC1/SC29/WG11 NO686, MPEG94/063.

ANNEX C

Technical Equipment and Listening Room Conditions

This report is divided into two sections, the first pertaining to the selection panel round of listening tests conducted at the Snell Technology Center in Valley Village, CA May 1-3, 1995, and the second pertaining to the expert listeners panel tests conducted at the National Cable Television Association screening room in Washington, DC May 15-17, 1995.

Listening Room Conditions for Selection Round

The Snell Technology Center is housed in an up-scale suburban home in a relatively quiet neighborhood in the San Fernando Valley, Los Angeles County, approximately 1 mile from a major freeway. The listening room employed for the selection round of the listening tests comprises a major part of the home, being 20'7" wide by 37' long by 9'2" high. With various intrusions and nooks the room volume is 6970 ft³. The room has had extensive acoustic treatment. The background noise, in the state in which the room was used for these tests, was measured approximately one week before the testing to be as follows:

Background Noise by Octave Band in dB re 20 μ Pa									
31.5 Hz	63 Hz	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz
46 dB	55 dB	44 dB	31 dB	23 dB	19 dB	19 dB	18 dB	20 dB	NA

Note: 20 μ Pa = 20 μ N/m²

Expressed as values obtained from Noise Criteria Curve analysis (the most widely used method for rating background noise in rooms in the U.S.) the following table ranks these measured sound pressure levels by octave band as NC values.

Background Noise by Octave Band expressed as NC values									
31.5 Hz	63 Hz	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz
NA	26	24	16	15	16	19	21	23	NA

On a wideband measurement, the sound pressure level is 30.3 dB, A weighted, slow reading, re 20 μ Pa. As a practical matter, the room meets the noise criteria curve NC-25. While this is somewhat higher than desirable for the most critical listening conditions, it is typical of suburban homes in such areas. Ameliorating the noise criteria rating is the fact that the noise spectrum is dominated by low frequency noise in the 63- and 125-Hz bands mostly originating from the freeway 1 mile away. Higher frequency bands meet NC-20 generally, with the high-frequency rise seen in the data being the probable limits of the instrumentation employed, as no high-frequency noise source could be identified by ear in the space.

The reverberation time by octave band was as follows:

Reverberation Time RT60 by Octave Band									
31.5 Hz	63 Hz	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz
2.43 s	0.71 s	0.36 s	0.22 s	0.20 s	*	0.17 s	0.18 s	0.17 s	0.16 s

* The measurement in this band was faulty, but there is no reason to believe that the 1 kHz band time would not lie between 0.15 and 0.2 s.

Note that the high frequency reverberation times are at the limit of the instrumentation in use, in particular, the decay time of the filters used in the analyzer employed.

These reverberation times are low for the room volume, yielding a higher ratio of direct sound to reflections and reverberation than is typical for a room of this volume. For the purposes of this test though, delivering such a high ratio of direct-to-reflected sound may well make the perception of transient events better than in typical spaces.

Control of specific early reflections was accomplished in this space by a combination of loudspeaker directivity, and room acoustic treatments. In particular, attention was paid to those surfaces where first-order reflections from the loudspeakers would reach the listening area.

Equipment used for Selection Round

The source tapes were delivered in the ADAT digital audio multi-track format, using 6 of the 8 available tracks. Two Alesis ADAT machines were used for playback, with the BRC accessory controller used to choose between the tapes (optically interconnected machines employing one set of output DACs for either source tape). The output signals from the tape under consideration was passed to a modified Madrigal Model PAV controller (modified for discrete multi-channel input). The controller extracted the low-bass portion of the program material from all the channels, summing them together and supplying the signal to the subwoofers, and provided a level matching function. The outputs of the PAV fed a DSP-based multi-channel equalizer, Snell Model RCS-1000 Room Correction System, consisting of one Motorola 56002 per channel with appropriate equalizing software. The equalizer outputs fed multiple NAD Model 208 power amplifiers for left, center, right main channels and subwoofers, and two NAD Model 2400 power amplifiers for the surrounds (bridged). The power amplifiers drove a Snell Music and Cinema Reference System, consisting of three front main loudspeaker channels, two surround loudspeaker channels, and two subwoofers.

The front loudspeakers are flat measured in a free field and in a listening window composed of 0° (axial), ±15° and ±30° horizontally, and ±15° vertically ±2 dB from the low-frequency crossover (80 Hz) to 20 kHz. The directivity index is 0 dB at low frequencies, and gradually increases to 7 dB ±2 dB over the range from 500 Hz to 10 kHz, then increases. There are no abrupt changes in the directivity index with respect to frequency, and the DI is achieved in such a way that the radiation pattern does not rotate from one plane to another across frequency. At 90 dB SPL and $f < 250$ Hz, no harmonic distortion component exceeds -30 dB, and for $f \geq 250$ Hz, -40 dB. The decay time for an abruptly stopped sine wave to 37% of the

steady-state level is $t_s < 5/f$, where f is the frequency. The group delay distortion is under 1 ms from 200 Hz to 8 kHz. The subwoofers are flat to within ± 2 dB from 17 Hz to 80 Hz, and have high level-playing capacity. The power response of the surround loudspeakers is flat to within ± 2 dB from the low-frequency crossover to 20 kHz. The equalization is set for a broad and smooth curve using two factors, known anechoic response of the loudspeakers, and spatially and temporally averaged response in the listening room. The response measured with a B&K type 4133 microphone averaged over the listening area is flat within ± 2 dB from 100 Hz to 10 kHz, and a gradual roll off above 10 kHz.

The placement of listeners in two rows, with the second row slightly elevated, provided direct line of sight from each of the loudspeakers to each of the listener positions. Nevertheless, listener locations were rotated so that each listener could spend time in each listening location. The subtended angles employed from the main listening locations were slightly narrower than recommended elsewhere, but sample listening at positions as specified by SMPTE ($\pm 30^\circ$ left and right front, $\pm 120^\circ$ for left and right surround) showed no difference in the results.

The Expert Listeners Panel tests were conducted in the screening room of the National Cable Television Association in Washington, DC.

Listening Room Conditions for Expert Listeners Round

The screening room of the National Cable Television Association is located in the NCTA building in Washington, DC. This space is a well-equipped screening room with both film and television projectors housed in an isolated projection booth, a motion-picture screen, and a multi-way sound system. The dimensions of the room are 21' wide by 44' long by approximately 18' high. An overhang of the projection booth into the space and other architectural features make the room volume 14,690 ft³.

Background noise testing was performed January 20, 1995 by Acoustical Design Collaborative, Inc. Further acoustic testing and sound system alignment was performed May 2, 1995 by Cardinal Sound. The background noise, in the state in which the room was used for these tests, was measured to be as follows:

Background Noise by Octave Band in dB re 20 μ Pa									
31.5 Hz	63 Hz	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz
52.5 dB	45.6 dB	35.4 dB	21.3 dB	15.6 dB	10.5 dB	9.3 dB	9.4 dB	7.7 dB	NA

The January measurements were taken with inherently lower noise test equipment, and so are those reported. The May measurements generally confirm the low-frequency noise measurements found in January (within 1–2 dB), but the higher frequency bands measure a higher noise level due to inherent self noise in the analyzer used in May (it is employed principally for motion-picture theater alignment and must be accurate to NC-30; the cause is the small microphones used to keep diffraction frequency response effects small).

Expressed as values obtained from Noise Criteria Curve analysis (the most widely used method for rating background noise in rooms in the U.S.) the following table ranks these measured sound pressure levels by octave band as NC values.

Background Noise by Octave Band expressed as NC values									
31.5 Hz	63 Hz	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz
NA	15	15	<<15	<<15	<<15	<<15	<<15	<<15	NA

The room meets the noise criteria curve NC-15 (it also meets the European standard NR-15).

The reverberation time by octave band was as follows:

Reverberation Time RT60 by Octave Band									
31.5 Hz	63 Hz	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	16 kHz
1.65 s	0.85 s	0.61 s	0.42 s	0.36 s	0.34 s	0.32 s	0.38 s	0.31 s	0.22 s

The reverberation time is well within industry standards for screening rooms of this volume, representing well similar spaces of this volume built for critical listening.

Control of discrete reflections is accomplished in this space through a combination of loudspeaker directivity and the design of specific room elements. For instance, the rear wall is made absorptive to prevent discrete screen loudspeaker sound reflections from arriving at listening locations with noticeable delayed reflections.

Equipment used for Expert Listening Panel tests

The tapes and the playback method were the same as in the selection panel round. The switched outputs of the ADAT machine were supplied to a Studio Interface adapter SI-1 accessory of a Dolby Model CP-200 Cinema Processor. The SI-1 interface adapter provides an input for multi-channel source audio, which is subsequently balanced for channel-to-channel level, volume controlled, and 1/3-octave band equalized for the main channels and parametrically equalized for the subwoofer and surround channels. The output of the CP-200 drives the input to the Apogee Model MPTS-1 three-way electronic crossover, which is custom tailored for use with the speaker complement employed. The crossover outputs drive multiple professional power amplifiers, QSC Model 1400. The main screen loudspeaker systems consist of three channels, each of which is tri-amplified, and a subwoofer. The screen channel system is an Apogee Model MPTS-1. Each surround array, left and right, consists of four JBL model 8330 loudspeakers.

Measurements made May 2 of frequency response of the three front channel systems show conformance to the standards SMPTE 202M-1991 and ISO 2969 Curve X within ± 2 dB from

40 Hz to 12.5 kHz and ± 3 dB from 31.5 Hz to 16 kHz as a spatial and temporal average over the listening area, with gradual rolloff at the frequency extremes. The "X curve" is the standard electro-acoustic response curve for motion-pictures and for that television dubbing which is done in conventional dubbing stages. An added equalizer in the Apogee MPTS-1 was activated, bringing the hf response up to better align with the response found in smaller room listening to flat axial-response loudspeakers as found in television listening. The resulting combination of the two curves, the X curve, and the hf equalizer, is an acoustical response which is down 3 dB from mid-range at 10 kHz. The subwoofer channel measured within ± 2 dB from 25 Hz to 125 Hz. The left and right surround response measures within ± 6 dB of the standards from 50 Hz to 16 kHz.

The seating area provided clear sight lines from each loudspeaker to each listener.

Page V-24

ANNEX D

Instructions to Subjects

1.0 General

The Federal Communications Commission (FCC) Advisory Committee on Advanced Television Standards is conducting formal listening tests of the audio quality of the audio coding system chosen for use in the U.S. HDTV standard. The multichannel audio system under test has five channels with three front channels/loudspeakers, L, R, C and two surround channels/loudspeakers, LS and RS. The stereo system uses L and R channels/loudspeakers.

The tests will be carried out without pictures.

The goals of the tests are:

- ~ to check the current audio quality reached with the coding algorithm.
- ~ to compare the quality of a previous system and the present system.

2.0 Test Procedure

The test procedure is taken from the ITU-R Recommendation "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems".

The tests have two phases: first a training phase involving groups of test subjects and the grading phase where subjects score the actual listening-test material.

2.1 Training Phase

The purpose of the training phase is to allow listeners to identify and become familiar with potential distortions and artifacts produced by the system under test. You will also become familiar with the test procedure. During this time you can and should comment on the items and discuss the artifacts heard with each other. After this training you should know "what to listen for."

We have an hour for this training phase which will include an opportunity to hear five of the ten test items which will be used in the test phase.

Although an exchange of views is expected during this training session, it is important that you should not discuss with the other listeners the grade that you, as an individual, would award, as this is a personal interpretation of the severity of the artifacts heard.

2.2 Test Phase

The test phase will be carried out in eight test sessions each lasting about 30 minutes. In each trial, you will hear three versions labeled "Ref," "A" and "B". "Ref" is always